EXHIBIT A

Fischer sequence
IVALPXGMLK

SEQ ID NO: 2 (single letter sequence)
YFPPPAAKEDFLGCLVKEIPPRLLYAKSSPAYPSVLGQTIRNSRWSSPDNVKPIYIVTPTNASHIQSAVVC
GRRHGVRIRVRSGGHDYEGLSYRSLQPEEFAVVDLSKMRAVWVDGKARTAWVDSGAQLGELYYAIHKASTV
LAFPAGVCPTIGVGGNFAGGGFGMLLRKYGIAAENVIDVKLVDANGTLHDKKSMGDDHFWAVRGGGGESFG
IVVAWKVRLLPVPPTVTVFKIPKKASEGAVDIINRWQVVAPQLPDDLMIRVIAQGPTATFEAMYLGTCQTL
TPMMSSKFPELGMNASHCNEMSWIQSIPFVHLGHRDNIEDDLLNRNNTFKPFAEYKSDYVYEPFPKRVWEQ
IFSTWLLKPGAGIMIFDPYGATISATPEWATPFPHRKGVLFNIQYVNYWFAPGAGAAPLSWSKEIYNYMEP
YVSKNPRQAYANYRDIDLGRNEVVNDVSTFSSGLVWGQKYFKGNFQRLAITKGKVDPTDYFRNEQSIPPLI
KKY

SIM+LALNVIEW analysis
37.5% identity in 8 residues overlap; Score: 22.0; Gap frequency: 0.0%

```
Fischer,        1 IVALPXGM
Phlp4(#2),    142 VLAFPAGV
                  * * *
```

SEQ ID NO: 4 (single letter sequence)
YFPPPAAKEDFLGCLVKEIPPRLLYAKSSPAYPSVLGQTIRNSRWSSPDNVKPIYIVTPTNASHIQSAVVC
GRRHGVRIRVRSGGHDYEGLSYRSLQPEEFAVVDLSKMRAVWVDGKARTAWVDSGAQLGELYYAIHKASPV
LAFPAGVCPTIGVGGNFAGGGFGMLLRKYGIAAENVIDVKLVDANGTLHDKKSMGDDHFWAVRGGGGESFG
IVVAWKVRLLPVPPTVTVFKIPKKASEGAVDIINRWQVVAPQLPDDLMIRVIAQGPTATFEAMYLGTCQTL
TPMMSSKFPELGMNASHCNEMSWIQSIPFVHLGHRDNIEDDLLNRNNTFKPFAEYKSDYVYEPFPKEVWEQ
IFSTWLLKPGAGIMIFDPYGATISATPEWATPFPHRKGVLFNIQYVNYWFAPGAGAAPLSWSKEIYNYMEP
YVSKNPRQAYANYRDIDLGRNEVVNDVSTFSSGLVWGQKYFKGNFQRLAITKGKVDPTDYFRNEQSIPPLI
KKY

SIM+LALNVIEW analysis
37.5% identity in 8 residues overlap; Score: 22.0; Gap frequency: 0.0%

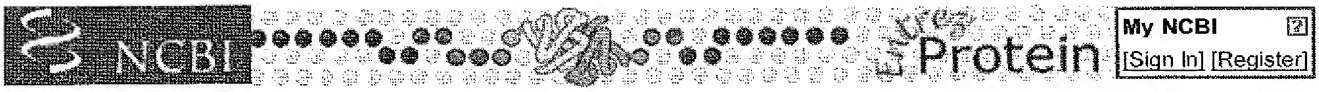
```
Fischer,        1 IVALPXGM
Phlp4(#4),    142 VLAFPAGV
                  * * *
```

SEQ ID NO: 6 (single letter sequence)
YFPPPAAKEDFLGCLVKEIPPRLLYAKSSPAYPSVLGQTIRNSRWSSPDNVKPLYIITPTNVSHIQSAVVC
GRRHSVRIRVRSGGHDYEGLSYRSLQPETFAVVDLNKMRAVWVDGKARTAWVDSGAQLGELYYAIYKASPT
LAFPAGVCPTIGVGGNFAGGGFGMLLRKYGIAAENVIDVKLVDANGKLHDKKSMGDDHFWAVRGGGGESFG
IVVAWQVKLLPVPPTVTIFKISKTVSEGAVDIINKWQVVAPQLPADLMIRIIAQGPKATFEAMYLGTCKTL
TPLMSSKFPELGMNPSHCNEMSWIQSIPFVHLGHRDALEDDLLNRNNSFKPFAEYKSDYVYQPFPKTVWEQ
ILNTWLVKPGAGIMIFDPYGATISATPESATPFPHRKGVLFNIQYVNYWFAPGAAAAPLSWSKDIYNYMEP
YVSKNPRQAYANYRDIDLGRNEVVNDVSTYASGKVWGQKYFKGNFERLAITKGKVDPTDYFRNEQSIPPLI
KKY

SIM+LALNVIEW analysis
42.9% identity in 7 residues overlap; Score: 19.0; Gap frequency: 0.0%

```
Fischer,        2 VALPXGM
Phlp4(#6),    143 LAFPAGV
                  * * *
```

EXHIBIT B



Item 1 - 4 of 4

1: P43213. Reports RecName: Full=Pol...[gi:1171008] BLink, Conserved Domains, Links

Next sequence

```
>gi|1171008|sp|P43213.1|MPAP1_PHLPR RecName: Full=Pollen
allergen Phl p 1; AltName: Full=Allergen Phl p I; AltName:
Allergen=Phl p 1; Flags: Precursor
MASSSSVLLVVVLFAVFLGSAYGIPKVPPGPNITATYGDKWLDAKSTWYGKPTGAGPKDNGGACGYKDVD
KPPFSGMTGCGNTPIFKSGRGCGSCFEIKCTKPEACSGEPVVVHITDDNEEPIAPYHFDLSGHAFGAMAK
KGDEQKLRSAGELELQFRRVKCKYPEGTKVTFHVEKGSNPNYLALLVKYVNGDGDVVAVDIKEKGKDKWI
ELKESWGAIWRIDTPDKLTGPFTVRYTTEGGTKTEAEDVIPEGWKADTSYESK
```

2: CAA81613. Reports pollen allergen P...[gi:3901094] BLink, Conserved Domains, Links

Previous sequence Next sequence

```
>gi|3901094|emb|CAA81613.1| pollen allergen Phl pI [Phleum
pratense]
MASSSSVLLVVALFAVFLGSAHGIPKVPPGPNITATYGDKWLDAKSTWYGKPTAAGPKDNGGACGYKDVD
KPPFSGMTGCGNTPIFKSGRGCGSCFEIKCTKPEACSGEPVVVHITDDNEEPIAAYHFDLSGIAFGSMAK
KGDEQKLRSAGEVEIQFRRVKCKYPEGTKVTFHVEKGSNPNYLALLVKFVAGDGDVVAVDIKEKGKDKWI
ALKESWGAIWRIDTPEVLKGPFTVRYTTEGGTKGEAKDVIPEGWKADTAYESK
```

3: 2118271A. Reports allergen Phl p I...[gi:1582250] BLink, Conserved Domains, Links

Previous sequence Next sequence

```
>gi|1582250|prf||2118271A allergen PhI p I
MASSSSVLLVVALFAVFLGSAHGIPKVPPGPNITATYGDKWLDAKSTWYGKPTAAGPKDNGGACGYKDVD
KPPFSGMTGCGNTPIFKSGRGCGSCFEIKCTKPEACSGEPVVVHITDDNEEPIAAYHFDLSGIAFGSMAK
KGDEQKLRSAGEVEIQFRRVKCKYPEGTKVTFHVEKGSNPNYLALLVKFSGDGDVVAVDIKEKGKDKWIA
LKESWGAIWRIDTPEVLKGPFTVRYTTEGGTKARAKDVIPEGWKADTAYESK
```

4: CAA55390. Reports Phl p I allergen ...[gi:473360] BLink, Conserved Domains, Links

Previous sequence

```
>gi|473360|emb|CAA55390.1| Phl p I allergen [Phleum
pratense]
MASSSSVLLVVVLFAVFLGSAYGIPKVPPGPNITATYGDKWLDAKSTWYGKPTGAGPKDNGGACGYKDVD
KPPFSGMTGCGNTPIFKSGRGCGSCFEIKCTKPEACSGEPVVVHITDDNEEPIAPYHFDLSGHAFGAMAK
KGDEQKLRSAGELELQFRRVKCKYPEGTKVTFHVEKGSNPNYLALLVKYVNGDGDVVAVDIKEKGKDKWI
ELKESWGAIWRIDTPDKLTGPFTVRYTTEGGTKTEAEDVIPEGWKADTSYESK
```

Help
General Help
Formats
Gaps
Matrix
References
ClustalW2 Help
ClustalW2 FAQ
Jalview Help
Scores Table
Alignment
Guide Tree
Colours

ClustalW2 Results

Results of search	
Number of sequences	4
Alignment score	9659
Sequence format	Pearson
Sequence type	aa
JalView	Launch
Output file	clustalw2-20091002-1639515291.output
Alignment file	clustalw2-20091002-1639515291.aln
Guide tree file	clustalw2-20091002-1639515291.dnd
Your input file	clustalw2-20091002-1639515291.input

SUBMIT ANOTHER JOB

To save a result file right-click the file link in the above table and choose "Save Target As".
If you cannot see the JalView button, reload the page and check your browser settings to enable Java Applets.

Scores Table

Sort by | Sequence Number | View Output File

SeqA	Name	Len(aa)	SeqB	Name	Len(aa)	Score
1	gi|1171008|	263	2	gi|3901094|	263	93
1	gi|1171008|	263	3	gi|1582250|	262	93
1	gi|1171008|	263	4	gi|473360|	263	100
2	gi|3901094|	263	3	gi|1582250|	262	98
2	gi|3901094|	263	4	gi|473360|	263	93
3	gi|1582250|	262	4	gi|473360|	263	93

PLEASE NOTE: Some scores may be missing from the above table if the alignment was done using multiple CPU mode. Please check the output.

Sort by | Sequence Number | View Output File

Alignment

Show Colors | View Alignment File

```
CLUSTAL 2.0.12 multiple sequence alignment


gi|1171008|       MASSSSVLLVVVLFAVFLGSAYGIPKVPPGPNITATYGDKWLDAKSTWYGKPTGAGPKDN 60
gi|473360|        MASSSSVLLVVVLFAVFLGSAYGIPKVPPGPNITATYGDKWLDAKSTWYGKPTGAGPKDN 60
gi|3901094|       MASSSSVLLVVALFAVFLGSAHGIPKVPPGPNITATYGDKWLDAKSTWYGKPTAAGPKDN 60
gi|1582250|       MASSSSVLLVVALFAVFLGSAHGIPKVPPGPNITATYGDKWLDAKSTWYGKPTAAGPKDN 60
                  ***********.*********:*****************************.******


gi|1171008|       GGACGYKDVDKPPFSGMTGCGNTPIFKSGRGCGSCFEIKCTKPEACSGEPVVVHITDDNE 120
gi|473360|        GGACGYKDVDKPPFSGMTGCGNTPIFKSGRGCGSCFEIKCTKPEACSGEPVVVHITDDNE 120
gi|3901094|       GGACGYKDVDKPPFSGMTGCGNTPIFKSGRGCGSCFEIKCTKPEACSGEPVVVHITDDNE 120
gi|1582250|       GGACGYKDVDKPPFSGMTGCGNTPIFKSGRGCGSCFEIKCTKPEACSGEPVVVHITDDNE 120
                  ************************************************************


gi|1171008|       EPIAPYHFDLSGHAFGAMAKKGDEQKLRSAGELELQFRRVKCKYPEGTKVTFHVEKGSNP 180
gi|473360|        EPIAPYHFDLSGHAFGAMAKKGDEQKLRSAGELELQFRRVKCKYPEGTKVTFHVEKGSNP 180
gi|3901094|       EPIAAYHFDLSGIAFGSMAKKGDEQKLRSAGEVEIQFRRVKCKYPEGTKVTFHVEKGSNP 180
gi|1582250|       EPIAAYHFDLSGIAFGSMAKKGDEQKLRSAGEVEIQFRRVKCKYPEGTKVTFHVEKGSNP 180
                  ****.******* ***:****************:*:**********************


gi|1171008|       NYLALLVKYVNGDGDVVAVDIKEKGKDKWIELKESWGAIWRIDTPDKLTGPFTVRYTTEG 240
gi|473360|        NYLALLVKYVNGDGDVVAVDIKEKGKDKWIELKESWGAIWRIDTPDKLTGPFTVRYTTEG 240
gi|3901094|       NYLALLVKFVACDGDVVAVDIKEKGKDKWIALKESWGAIWRIDTPEVLKGPFTVRYTTEG 240
gi|1582250|       NYLALLVKFS-GDGDVVAVDIKEKGKDKWIALKESWGAIWRIDTPEVLKGPFTVRYTTEG 239
                  ********:  ********************* ***************: *.***********


gi|1171008|       GTKTEAEDVIPEGWKADTSYESK 263
gi|473360|        GTKTEAEDVIPEGWKADTSYESK 263
gi|3901094|       GTKGEAKDVIPEGWKADTAYESK 263
gi|1582250|       GTKARAKDVIPEGWKADTAYESK 262
                  *** .*:***********:****
```

PLEASE NOTE: Showing colors on large alignments is slow

Show Colors | View Alignment File

Guide Tree

Show as Phylogram Tree | Show Distances | View DND File

```
(
gi|1171008|:0.00000,
(
gi|3901094|:0.00369,
gi|1582250|:0.00776)
:0.06095,
gi|473360|:0.00000);
```

Unveiling the secrets of the primary structure of Phl p 4

Molecular cloning of the major pollen allergen from Timothy Grass (*Phleum pratense*)

A. Nandy, S. Buchhop, R. Suck, A. Petersen*, O. Cromwell, H. Fiebig

Allergopharma Joachim Ganzer KG, R&D Department, 21465 Reinbek, Germany
*Research Center Borstel, Biochemical & Molecular Allergology, Borstel, Germany
Contact e-mail: andreas.nandy@allergopharma.de

Introduction

Grass pollen allergy is one of the most common allergies worldwide. Recombinant allergens are believed to represent the future of allergen specific immunotherapy. Whereas the cDNA sequences of several grass pollen allergens are known, the coding sequence for Phl p 4, a major grass pollen allergen recognised by more than 70 % of allergic patients (1-5), has so far escaped detection (5).

Methods

A set of degenerate oligonucleotide primers was designed based on N-terminal and internal protein sequences obtained from purified natural Phl p 4 (Tab. 1). In a complex PCR strategy (Fig. 1) involving degenerate and specific primers the Phl p 4 gene could be amplified from genomic DNA and from cDNA derived from *Phleum pratense* pollen.

Results

The deduced amino acid sequence of full length Phl p 4 contains 500 amino acids, with a calculated MW of 55,7 kDa and a calculated basic pI of 8,8 (Tab. 2). The identity of the Phl p 4 sequence has been confirmed by positive reaction of recombinant Phl p 4 with specific monoclonal antibodies (Fig. 2) and by reaction with IgE from grass pollen allergics (Fig. 3). A sequence database homology search revealed similarities to a group of berberine bridge enzyme-like oxido-reductases (Fig. 4).

Tab. 2 Phl p 4 Sequence analysis

Amino acid		Number	% by weight	% by frequency
A	Ala	40	5.50	8.00
C	Cys	5	[illegible]	1.00
D	Asp	24	4.95	4.8
E	Glu	22	5.10	4.40
F	Phe	24	6.34	4.80
G	Gly	42	4.30	8.40
H	His	10	2.46	2.00
I	Ile	29	5.92	5.80
K	Lys	29	4.67	5.80
L	Leu	33	6.70	6.60
M	Met	11	2.57	2.20
N	Asn	22	4.40	4.40
P	Pro	39	6.80	7.80
Q	Gln	15	3.45	3.00
R	Arg	24	6.73	4.80
S	Ser	32	5.01	6.40
T	Thr	21	3.81	4.20
V	Val	41	7.29	8.20
W	Trp	13	4.34	2.60
Y	Tyr	24	7.02	4.80

Sequence analysis	Calculated values
Molecular weight	55,734 Da
Length	500 amino acids
Isoelectric point	8.80
N-Glycosylation sites	4
FAD binding consensus sequence	Yes *

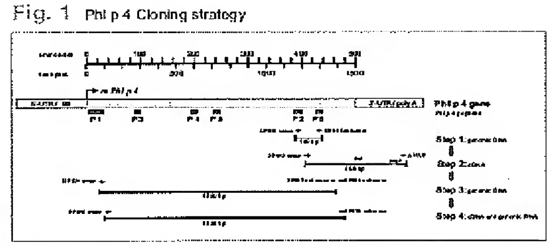
* To date the binding of a flavin co-factor could not be proved for purified natural or recombinant Phl p 4.

Tab. 1 N-terminal and internal peptide sequences of Phl p 4

Peptide	Sequence	From	To	Fragment
P 1 a	YFPPP AAKED FLGXL VKE P PRLLY AKSSP AYP	1	33	N-terminal
P 1 b	YFPPP AAKED FLGXL VKE P PRLLY AKSSP	1	30	N-terminal
P 1 c	YFPXX AAKED FLGXL	1	15	N-terminal
P 1 d	YFPXX AAKED FLGXL	1	15	N-terminal
P 1 e	YFPXX AAKDD FLGXL	1	15	N-terminal
P 1 f	YFPXX LANED F	1	11	N-terminal
P 2	SAIPP XARKG VLFNI QYV	384	401	Glu C
P 3	GLXYR XLDPE	90	99	Glu C
P 4	KXXGD DHFXA VR	194	205	Lys C
P 5	APEGA VDI	238	245	Lys C
P 6	MBPYV SNPV GAYAN Y	425	439	CNBr

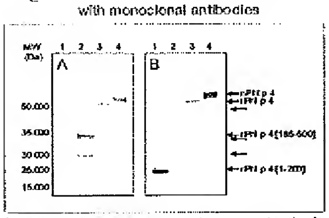
N-terminal sequencing of purified natural Phl p 4 and of fragments obtained from protease digestion or CNBr cleavage revealed the peptide sequences P1-P6. P1 a-f presumably represent different variants of mature Phl p 4.

Fig. 1 Phl p 4 Cloning strategy



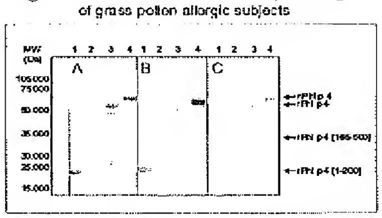
Step 1: Degenerate oligonucleotide primer pools (DP#30 and DP#37) based on the Phl p 4 peptide sequences P2 and P6 have been used in a PCR reaction to amplify a small 149 bp internal DNA fragment of genomic *Phleum pratense* DNA.
Step 2: Based on that sequence a specific oligonucleotide primer (SP#82) was designed to be used in a 3'-RACE PCR approach in combination with the anchor primer AUAP (Life technologies). A 489 bp fragment spanning the entire 3'-end of the Phl p 4 gene could be obtained.
Step 3: Two specific antisense primers (SP#85 and SP#86) have been designed on the basis of that sequence. These were used in combination with a degenerate sense primer pool (DP#29) based on the N-terminal peptide P1 to amplify genomic Phl p 4 DNA in a "sense side mixed oligonucleotide, antisense side nested unique primer" PCR approach. A resulting 1320 bp fragment has been cloned and sequenced. The sequences of the peptides P2-P6 could be aligned with the deduced amino acid sequence of this fragment confirming the identity with the Phl p 4 gene.
Step 4: A specific sense primer (SP#88) has been designed and was used together with SP#86 antisense to amplify a 1336 bp fragment of Phl p 4. Several independent PCR-products from genomic DNA as well as from cDNA have been sequenced to exclude PCR errors. Different variants of Phl p 4 could be detected. The identity of cDNA and genomic clones show that no intron sequences are present in the amplified region.

Fig. 2 Reaction of recombinant Phl p 4 with monoclonal antibodies



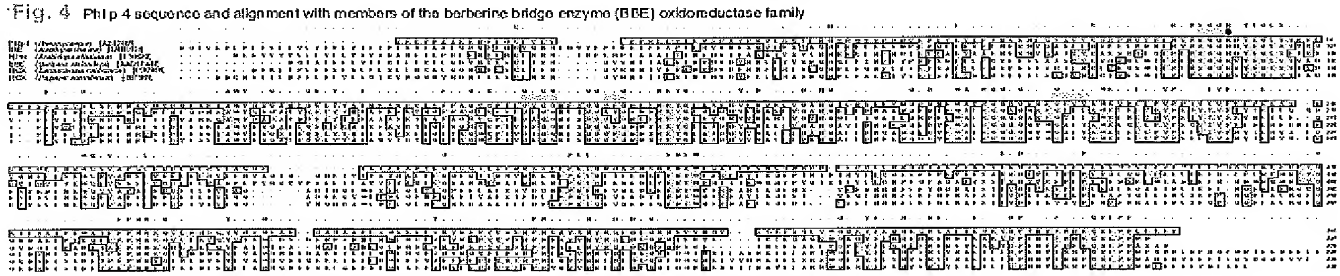
Western blot of whole cell extracts of E. coli expressing (1) a N-terminal fragment of Phl p 4 (aa 1-200, MW ~ 22 kDa), (2) a C-terminal fragment of Phl p 4 (aa 185-500, MW ~ 36 kDa, (3) full length recombinant Phl p 4, and (4) purified natural Phl p 4.
A: The monoclonal antibody 5H1 detects nPhl p 4, rPhl p 4 and the C-terminal fragment. B: The monoclonal antibody 3C4 detects nPhl p 4, rPhl p 4 and the N-terminal fragment of Phl p 4.

Fig. 3 Reaction of recombinant Phl p 4 with IgE of grass pollen allergic subjects



Western blot of whole cell extracts of E. coli expressing (1) a N-terminal fragment of Phl p 4 (aa 1-200, MW ~ 22 kDa), (2) a C-terminal fragment of Phl p 4 (aa 185-500, MW ~ 36 kDa, (3) full length recombinant Phl p 4, and (4) purified natural Phl p 4. Incubation with sera of three different grass pollen allergic individuals (A, B, C) confirmed the IgE reactivity of recombinant Phl p 4.

Fig. 4 Phl p 4 sequence and alignment with members of the berberine bridge enzyme (BBE) oxidoreductase family



Multiple alignment of Phl p 4, BBE (berberine bridge enzyme), HPR (hypothetical protein) and ROX (reticuline oxidase) sequences. Residues identical to the Phl p 4 sequence are shown in yellow. Phl p 4 shares 34 % identical amino acid residues with BBE from E. californica and 41 % with BBE from A. thaliana. The oxidoreductase family flavin-binding consensus sequences are marked with blue lines ( ), the conserved histidine to which the FAD is covalently bound in BBE is marked in red (●).

Conclusion

The ability to produce recombinant Phl p 4, a major allergen of grass pollen with one of the highest IgE binding frequencies measured in sera of pollen allergic patients, may represent a key step for the development of future diagnostic and immunotherapeutic preparations. Recombinant Phl p 4 will also serve as a valuable tool to elucidate the role of the carbohydrate moiety of natural Phl p 4 in IgE reactivity and cross-reactivity with other plant and food allergens.

References

1) R. Suck, S. Hagen, O. Cromwell, H. Fiebig (2000), Clin. Exp. Allergy, 30, 1395-1402
2) R.E. Rossi, G. Monasterolo, S. Monasterolo (2001), Allergy 56, 1180-11853
3) S. Stumvoll, J. Lidholm, R. Thunberg, A. DeWitt, P. Eibensteiner, I. Swoboda, A. Bugajska-Schretter, S. Spitzauer, L. Vangelista, L. Kazemi-Shirazi, W.A. Sperr, O. Kraft, R. Valenta (2002), Biol. Chem., 383, 1383-1396
4) A. Mari (2003), Clin. Exp. Allergy, 33, 43-51
5) K. Andersson, J. Lidholm (2003), Int. Arch. Allergy Immunol., Review article, 130, 87-107

First Presentation, June 2003, EAACI, Paris Contact: Allergopharma Joachim Ganzer KG D-21462 Reinbek, Germany Tel.: ++49/40/7 27 65-0 Fax ++49/40/7 22 77 13

DNA sequences of group 4 allergens from rye, wheat, barley and *Lolium perenne*

Comparison with isoforms of *Phleum pratense* Phl p 4

A. Nandy, M. Wald, L. Gräfe, O. Cromwell, H. Fiebig

Allergopharma Joachim Ganzer KG, R&D Department, 21465 Reinbek, Germany
Contact e-mail: andreas.nandy@allergopharma.de

Introduction

Grass pollen allergy is one of the most important allergic diseases world-wide. Several grass species grown in meadows, like *P. pratense* and *L. perenne*, contribute to allergic sensitisations, but also allergens from extensively cultured cereals, especially rye, make a profound contribution to the development of allergy. The group 4 major allergen of *P. pratense*, Phl p 4, is recognised by more than 70 % of grass allergic patients[1,2,3]. IgE-binding cross-reactivity has been described for some group 4 allergens of different grass species[3], but until now only the Phl p 4 gene could be deciphered on the DNA-level.

Results

The Pooideae group 4 allergens represent a family of basic proteins with molecular weights of about 55 kDa and calculated pI values far above 8 (Tab. 1, Fig. 1). In rye, wheat and *P. pratense* distinct isoforms with amino acid identities of 88 to 94 % could be detected. Additionally these isoforms exist in different minor variants. The inter-species homology lies in the range 83 % (Phl p 4 to Triticeae species) to 95 % (Sec c 4 to Tri a 4) (Fig. 2, Fig. 3).

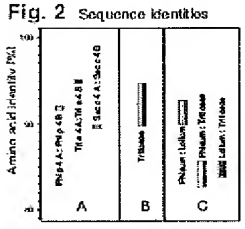
Tab. 1 Sequence analysis of grass pollen group 4 allergens

Protein	Source	Sequence length (amino acids)	Isoelectric point (pI)	Molecular weight (Da)
Phl p 4 A	*Phleum pratense*	500	8,8	55.895
Phl p 4 B	*Phleum pratense*	500	9,2	55.624
Lol p 4*	*Lolium perenne*	423 (fragment)	8,8*	*
Sec c 4 A	*Secale cereale* (rye)	498	9,1	54.930
Sec c 4 B	*Secale cereale* (rye)	498	9,3	54.903
Tri a 4 A	*Triticum aestivum* (wheat)	497	8,8	55.237
Tri a 4 B	*Triticum aestivum* (wheat)	497	8,8	55.149
Hor v 4	*Hordeum vulgare* (barley)	496	9,3	54.815

The sequence length, isoelectric points and molecular weight calculations were made on the basis of the mature proteins. For Phl p 4 the N-terminal residue has been determined by N-terminal protein sequencing. Based on the homology alignment (Fig. 1) the putative cleavage sites of Triticeae spp. have been used for calculation.

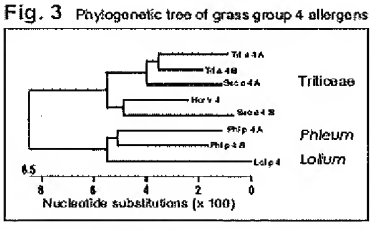
* The Lol p 4 sequence is only partial and contains about 85 % of the mature Lol p 4 sequence.

Fig. 2 Sequence identities



The amino acid identities were calculated on the basis of the mature allergens. In case of Lol p 4 the overlapping region has been used for calculation.
A The sequence identities of intra-species variants of group 4 allergens range from 89 % (Sec c 4) to 95 % (Tri a 4). The two major variants of Phl p 4 show intermediate identities of 92 %.
B The identities of allergens of the Triticeae species range from 90 % (Sec c 4 B to Tri a 4) to 95 % (Sec c 4A to Tri a 4 B).
C Inter-species identities of members of the genera *Phleum*, *Lolium* and of the 3 Triticeae genera *Hordeum*, *Triticum* and *Secale*.

Fig. 3 Phylogenetic tree of grass group 4 allergens

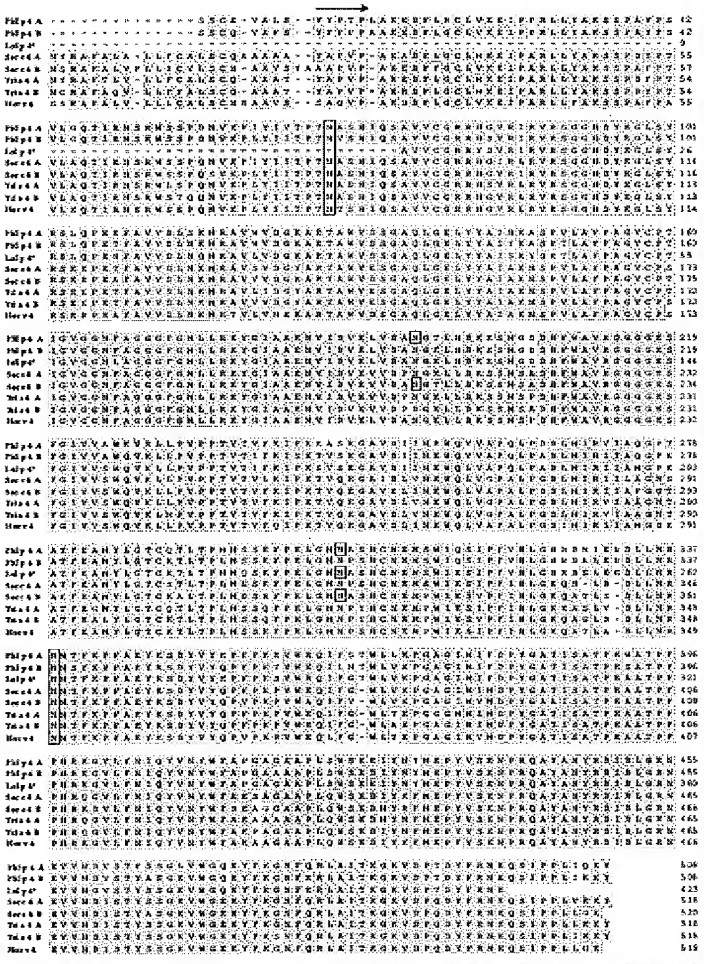


The dendogram[4] illustrates the phylogenetic relationships of the grass group 4 sequences. The rooted tree has been generated by using the DNA sub-sequences that overlap the Lol p 4 fragment (1272 bp).
Remarkably intra-species variants (eg. Sec c 4 A and Sec c 4 B) show sequence identities similar to those of sequences originating from different Triticeae species (compare also Fig. 2A and 2B). The same can be seen for *Phleum pratense* variants that have similar degrees of amino acid difference as compared to the *Lolium perenne* sequence (compare also Fig. 2A and 2C).

Methods

Based on the DNA sequence of Phl p 4 several PCR-primer sequences with cross-reactivity to DNA sequences of related species could be designed. The group 4 DNA sequences of *Lolium perenne* (Lol p 4), *Secale cereale* (Sec c 4), *Hordeum vulgare* (Hor v 4), and *Triticum aestivum* (Tri a 4) have been amplified, cloned and sequenced.

Fig. 1 Deduced amino acid sequence alignment of grass pollen group 4 allergens



Multiple alignment of *Phleum pratense* Phl p 4 variant forms, *Lolium perenne* Lol p 4, *Secale cereale* (rye) Sec c 4 variant forms, *Triticum aestivum* (wheat) Tri a 4 variant forms, and *Hordeum vulgare* (barley) Hor v 4. Residues that match the consensus sequence are shaded in yellow. The start of the mature Phl p 4 sequence as deduced by N-terminal protein sequencing of purified natural Phl p 4 is marked with a red arrow. Potential N-glycosylation sites are marked with blue boxes.

* The Lol p 4 sequence is only partial and contains about 85 % of the mature Lol p 4 sequence.

Conclusion

The group 4 allergens represent a family of proteins that are conserved among different grass species. The occurrence of cross-reacting isoforms in distinct species with amino acid homologies that are comparable to those of different group 4 molecules across the species border is remarkable. Since recombinant group 4 allergens may be important for a future recombinant allergen based specific immunotherapy, strong efforts should be made to evaluate the cross-reactive therapeutic potential of the different group 4 allergens and their isoforms.

References

1) R.E. Rossi et al. (2001), Allergy 56, 1180-11853
2) K. Andersson and J. Lidholm (2003), Int. Arch. Allergy Immunol., Review article, 130, 87-107
3) S. Stumvoll et al. (2002), Biol. Chem. 383, 1383-1396
4) Lasergene DNASTAR, Inc., Madison, WI 53715, U.S.A.

Recombinant *Phleum pratense* pollen allergen Phl p 4

Clues to new data for an old allergen?

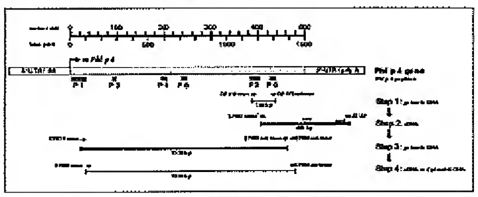
A. Nandy, M. Wald, B. Weber, H. Kahlert, O. Cromwell, H. Fiebig

Allergopharma Joachim Ganzer KG, R&D Department, 21465 Reinbek, Germany
Contact e-mail: andreas.nandy@allergopharma.de

Introduction

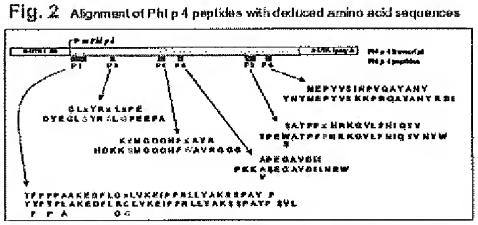
The group 4 allergens of grasses were first described more than 20 years ago and are well known as important major allergens of grass pollen allergy, one of the most common allergies world-wide. Phl p 4 is a basic glycoprotein that, together with Phl p 13, accounts for the high molecular weight fraction of grass pollen allergens. Frequencies of IgE sensitisation higher than 70% have often been reported (1-3), and therefore Phl p 4 seems to be as important as Phl p 5. Contrary to the situation for Phl p 5 and other important Phleum allergens, the primary structure of Phl p 4 has been discovered only recently, despite very considerable efforts in the past.

Fig. 1 Phl p 4 cloning strategy



Step 1: Degenerate oligonucleotide primer pools (DP#30 and DP#37) based on the Phl p 4 peptide sequences P2 and P6 have been used in a PCR reaction to amplify a small 149 bp internal DNA fragment of genomic *Phleum pratense* DNA.
Step 2: Based on that sequence a specific oligonucleotide primer (SP#32) was designed to be used in a 3'-RACE PCR approach in combination with the anchor primer AUAP (Life technologies). A 489 bp fragment spanning the entire 3'-end of the Phl p 4 gene could be obtained.
Step 3: Two specific antisense primers (SP#85 and SP#86) have been designed on the basis of that sequence. These were used in combination with a degenerate sense primer pool (DP#29) based on the N-terminal peptide P1 to amplify genomic Phl p 4 DNA in a "sense side mixed oligonucleotide, antisense side nested unique primer" PCR approach. A resulting 1320 bp fragment has been cloned and sequenced.
Step 4: A specific sense primer (SP#88) has been designed and was used together with SP#86 antisense to amplify a 1336 bp fragment of Phl p 4. Several independent PCR products from genomic DNA as well as from cDNA have been sequenced to exclude PCR errors. Different variants of Phl p 4 could be detected. The identity of cDNA and genomic clones show that no intronic sequences are present in the amplified region.

Fig. 2 Alignment of Phl p 4 peptides with deduced amino acid sequences

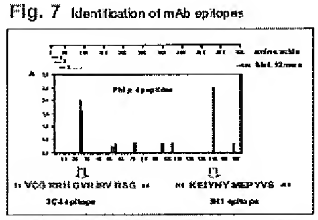


Natural Phl p 4 derived N-terminal and internal peptide sequences (first lines) served to confirm the deduced Phl p 4A (second lines) and Phl p 4B (third lines) genomic and cDNA sequences.

Results

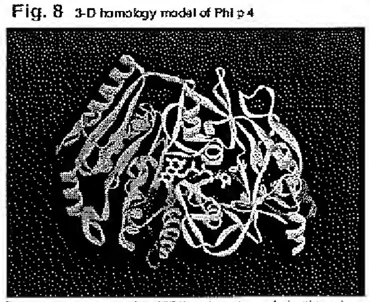
The experimental procedure that in the end led to the genomic and cDNA sequences of the gene was based on a complex PCR strategy involving specific and degenerate primers (Fig. 1). The identified sequence has been confirmed to be Phl p 4 by alignment of the deduced amino acid sequence with natural nPhl p 4 derived peptides (Fig. 2). The deduced amino acid sequences of two variants of mature Phl p 4 consist of 500 amino acids each, with calculated molecular weights of 56 kDa and basic pI's of 8,8 and 9,2, respectively. A sequence database homology search revealed similarities to berberine bridge enzyme-like oxido-reductases (Fig. 3). Recombinant Phl p 4 was expressed in *E. coli* as inclusion bodies and has been subjected to a refolding procedure. However, the correct folding turned out to be difficult to achieve. Therefore we have expressed Phl p 4 in the methylotrophic yeast *Pichia pastoris*. The *P. pastoris* derived Phl p 4 is highly soluble and has been purified via His-tag from culture supernatants. Purified recombinant Phl p 4 has been characterised by SDS-PAGE (Fig. 4), IgE inhibition assay (Fig. 5), and protein dots using monoclonal antibodies, as well as IgE containing allergic subjects' sera (Fig. 6). The epitopes of two monoclonal antibodies 3C4, and 5H1 could be localised to the N-terminal and C-terminal domain, respectively (Fig. 7). A 3-D model of Phl p 4 was generated on the basis of the vanillyl-alcohol oxidase (VAO) structure (Fig. 8).

Fig. 7 Identification of mAb epitopes



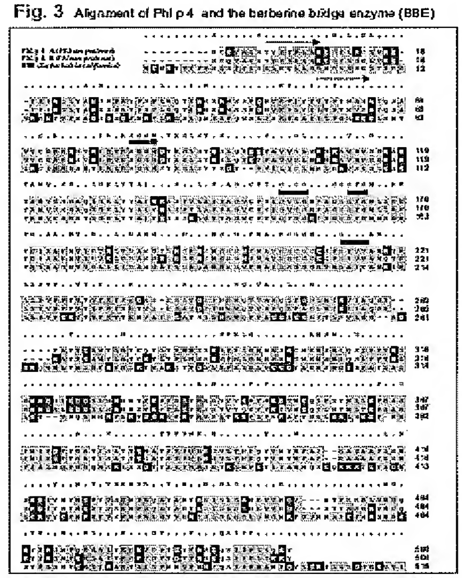
Overlapping 12mer peptides have been synthesized, biotinylated, and adsorbed to streptavidin coated MTP. The bound Phl p 4 specific monoclonal antibodies were detected by AP-α-mouse-IgG.

Fig. 8 3-D homology model of Phl p 4



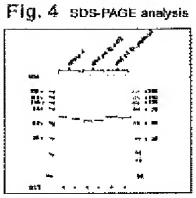
The consensus sequences of the BBE-like oxido-reductase family (5) have been used to align the Phl p 4 and vanillyl-alcohol oxidase (VAO) sequences. A 3-D homology model of Phl p 4 has been generated on the basis of the known structure of VAO (6) using the program Discovery Studio Modeling, Accelrys, St. Diego, U.S.A.

Fig. 3 Alignment of Phl p 4 and the berberine bridge enzyme (BBE)



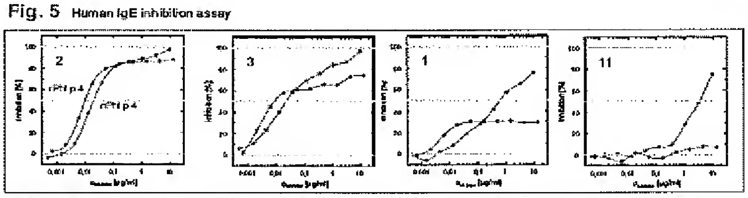
Sequence alignment of *Phleum pratense* Phl p 4 variant A (EMBL AJ512487), variant B (EMBL AJ512488) and *Eschscholzia californica* berberine bridge enzyme (SWISSPROT P30986). The consensus line shows identical residues. The start of the mature Phl p 4 sequence as deduced by N-terminal protein sequencing of purified natural Phl p 4, and the first amino acids of BBE are marked with red arrows. Cysteines are shaded in yellow, acidic residues in red, basic residues in blue, and hydrophobic residues are shown in green.
The oxido-reductase family flavin-binding consensus sequences are marked with black lines, the histidine that serves as the covalent attachment site for the FAD co-factor in BBE (4) is marked with a red dot.

Fig. 4 SDS-PAGE analysis



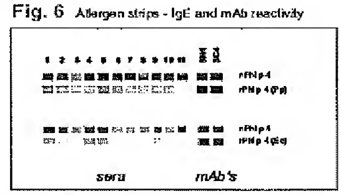
SDS-PAGE comparison of natural nPhl p 4, recombinant rPhl p 4 expressed in *E. coli*, and expressed in *P. pastoris*

Fig. 5 Human IgE inhibition assay



IgE inhibition assay using sera of grass pollen allergic blood donors and purified natural Phl p 4 as solid phase. The inhibitory capacity of natural Phl p 4 (green dots) and recombinant Phl p 4B-His purified from *P. pastoris* (red dots) expression cultures were compared. The sera of subjects 1, and 11 are known to contain significant amounts of cross reacting carbohydrate determinant (CCD) IgE antibodies.

Fig. 6 Allergen strips - IgE and mAb reactivity



Allergen strips comparing rPhl p 4B-His purified from *P. pastoris* expression cultures and rPhl p 4A expressed in *E. coli*. The *E. coli* derived protein has been purified from inclusion bodies and was solubilized in 4M GnHCl prior to dotting.

Conclusion

The ability to produce recombinant Phl p 4 may represent a key step for the development of future diagnostic and immunotherapeutic preparations and may be of special importance for those allergic persons that show a strong IgE response to Phl p 4.

References

1) R.E. Rossi et al. (2001), Allergy 56, 1180-1185
2) S. Stumvoll et al. (2002), Biol. Chem. 383, 1383-1396
3) A. Mari (2003), Clin. Exp. Allergy, 33, 43-51
4) T.M. Kutchan and H. Dittrich (1995), J. Biol. Chem. 270, 24475-24481
5) M.W. Fraaije et al. (1998), TIBS 23, 206-207
6) Mattevi et al. (1997), PDB ID: 1VAO